JCIM JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

Cite This: J. Chem. Inf. Model. 2018, 58, 982−992

pubs.acs.org/jcim

Peptidic Macrocycles - Conformational Sampling and Thermodynamic Characterization

Anna S. Kamenik,[†] Uta Lessel,[‡] Julian E. Fuchs,[§] Thomas Fox,*[,‡] and Klaus R. Liedl*[,†]

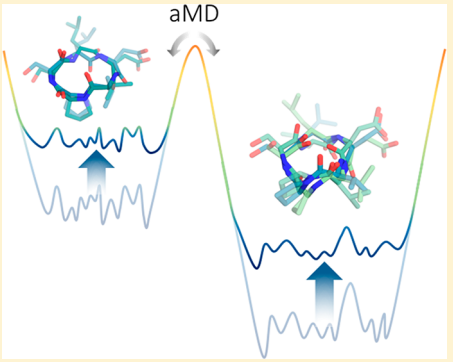
[†]Institute of General, Inorganic and Theoretical Chemistry, Center for Molecular Biosciences Innsbruck, University of Innsbruck, 6020 Innsbruck, Austria

[‡]Medicinal Chemistry, Boehringer Ingelheim Pharma GmbH & Co. KG, 88397 Biberach, Germany

[§]Department of Medicinal Chemistry, Boehringer Ingelheim RCV GmbH & Co KG, 1120 Vienna, Austria

Supporting Information

ABSTRACT: Macrocycles are of considerable interest as highly specific drug candidates, yet they challenge standard conformer generators with their large number of rotatable bonds and conformational restrictions. Here, we present a molecular dynamics-based routine that bypasses current limitations in conformational sampling and extensively profiles the free energy landscape of peptidic macrocycles in solution. We perform accelerated molecular dynamics simulations to capture a diverse conformational ensemble. By applying an energetic cutoff, followed by geometric clustering, we demonstrate the striking robustness and efficiency of the approach in identifying highly populated conformational states of cyclic peptides. The resulting structural and thermodynamic information is benchmarked against interproton distances from NMR experiments and conformational states identified by X-ray crystallography. Using three different model systems of varying size and flexibility, we show that the method reliably reproduces experimentally determined structural ensembles and is capable of identifying key conformational states that include the bioactive conformation. Thus, the described approach is a robust method to generate conformations of peptidic macrocycles and holds promise for structure-based drug design.



INTRODUCTION

Macrocycles are an intriguing compound class, as they promise to address long-standing druggability challenges such as protein−protein interfaces with remarkably high affinity.[1−3] Especially peptidic macrocycles show significantly increased activity and bioavailability compared to their acyclic counterparts.[4] The advantages of peptidic macrocycles were demonstrated for several challenging drug targets, such as the HIV-1 protease,[5,6] where the cyclization of peptidic compounds increased their affinity by up to 4 orders of magnitude. The striking affinity enhancement due to macrocyclization, which has been observed repeatedly, is proposed to originate from a structural preorganization:[6−8] In macrocyclic compounds, the ring closure reduces the accessible conformational states which ideally would be able stabilize the peptide in its bioactive conformations. Following the concept of conformational selection, this shift of state populations toward the active state favors binding.[9] However, in case the macrocyclization stabilizes a nonbioactive conformational state in solution, it could also slow down binding. Additionally, the conformational restraints decrease the loss in conformational entropy upon binding, which contributes to the increased affinities found for macrocyclic compounds.[10]

Hence, macrocycles are among the most promising compound classes in the "beyond-rule-of-5" orally available drug space.[4,11−13] Peptidic macrocycles in particular benefit from the conformational restraints, which allow them to bypass the fast proteolytic degradation observed for their acyclic analogs.[14] Peptidic macrocycles are typically characterized by a large polar surface area, which presumably hinders membrane permeability.[15] Nevertheless, it has already been demonstrated that the exceptional structural characteristics of peptidic macrocycles can allow considerably high permeability and bioavailability.[16,17] In the prominent case of cyclosporine A, the observed permeability is attributed to a particular N-methylation pattern which promotes a conformational switch upon changes in solvent polarity.[18−21] Several other studies on varying model systems also reported a remarkable permeability and metabolic stability for macrocyclic peptides with rigidified scaffolds.[6,22]

To perform rational drug design, a reliable and extensive description of the conformational space of putative ligands is vital.[23,24] For peptidic macrocycles, this task is particularly important as their benefits in selectivity and bioavailability originate from the exceptional characteristics of their structural ensembles.[25] Hence, a computational characterization of their conformational space may significantly aid the drug development process for macrocyclic compounds.

Received: February 19, 2018
Published: April 13, 2018

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

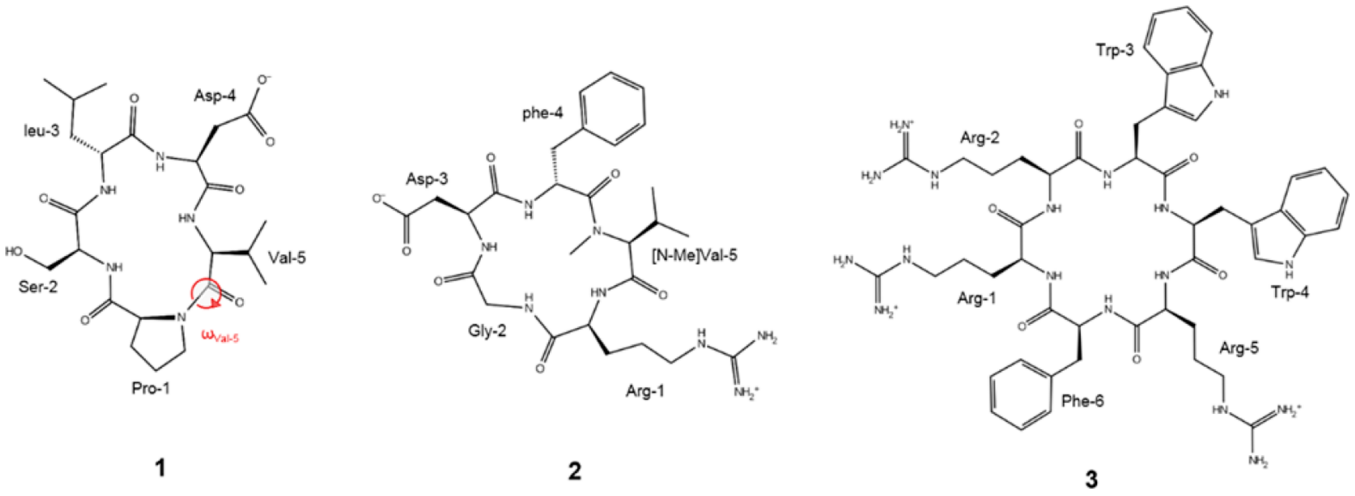


Figure 1. Studied peptidic macrocycles. (1) The integrin-binding pentapeptide cyclo-(Pro-Ser-leu-Asp-Val), where $\omega_{Val\text{-}5}$, (the peptide bond between Val-5 and Pro-1) distinguishes the *cis*- and *trans*-proline state, (2) the anticancer drug candidate cyclo-(Arg-Gly-Asp-phe-([N-Me]Val)), known as cilengitide, and (3) the antimicrobial hexapeptide cyclo-(Arg-Arg-Trp-Trp-Arg-Phe).

Several comprehensive studies demonstrated that standard algorithms to generate small molecule conformations fail to sample physically meaningful ensembles of macrocycles.[26,27] Hence, dedicated macrocycle conformation generators were developed, but some challenges yet remain.[14,28,29] While these state-of-the-art algorithms excel in creating a diverse ensemble of macrocycle conformations, they are typically not designed to provide reliable thermodynamic insights.[14,30,31] Peptidic macrocycles are particularly challenging even for dedicated macrocycle conformation generators due to their unique structural preferences, dynamics, and biophysical properties.[14,30,32] However, a recent study demonstrated the reliability of current force-field-based sampling techniques by elucidating several key features defining the conformational space of peptidic macrocycles.[32] Also other force-field-based studies applying enhanced sampling techniques, such as metadynamics[33,34] and replica-exchange MD,[35] to characterize the structural ensemble of cyclic peptides showed highly promising results.[36−38]

In the present work, we show accelerated molecular dynamics (aMD) simulations[39] to be a strikingly reliable tool to gather a diverse and thermodynamically meaningful conformational ensemble for peptidic macrocycles. In aMD simulations, the introduction of a bias potential decreases the height of the barrier between individual conformational states. Thus, aMD simulations sample the conformational space more efficiently than conventional MD (cMD) simulations, yet allow the recalibration to the "real" phase space potential energy surface (PES).[40−42] Two further advantages of aMD simulations are that no prior knowledge of the PES is required and that the computational effort is comparable to cMD simulations. The approach is well established as demonstrated in several studies where aMD simulations capture diverse conformational ensembles of biomolecules while preserving the thermodynamic information on the original PES.[43−47]

For this first evaluation of the proposed aMD-based approach, we exclusively considered data from NMR experiments in aqueous solution. An enhanced understanding of macrocycles in aqueous solution is urgently needed to enhance our understanding of conformational preorganization for binding. Despite the great interest in macrocyclic peptides as potent drug candidates, we found that the available experimental data on their structural ensembles in aqueous solution are rather sparse, probably due to their limited solubility in water.[48]

Additionally, organic solvents still hold numerous challenges in MD-based computational modeling. There is much less experience for MD simulations in explicit organic solvents compared to water, which results in a limited reliability of available force field parameters.[49−51] Moreover, the correlation times, which describe the time scale needed for molecular reorientation, are much longer for organic solvents than for water, and thus the necessary simulation time increases significantly.[52]

To keep the computational effort tractable, we further restricted the size of our model systems to a maximum of 12 rotatable bonds, corresponding to 6 amino acid residues. We profile the conformational space of three different cyclopeptidic model systems of varying size, flexibility, and available experimental data (Figure 1). The first system studied is the integrin-binding pentapeptide cyclo-(Pro-Ser-leu-Asp-Val)[53] for which NOE restraints in aqueous solution and the relative prevalence of two of its major components, a *cis*- and a *trans*-proline state, were reported.[54] Furthermore, we characterize the structural ensemble of the cyclic pentapeptide cilengitide, cyclo-(Arg-Gly-Asp-phe-[(N-Me)-Val]). Cilengitide has been studied as an anticancer drug candidate and its conformational space in aqueous solution as well as its target-bound structure have been thoroughly characterized using NMR and X-ray crystallography.[55−57] The third system is the antimicrobial hexapeptide cyclo-(Arg-Arg-Trp-Trp-Arg-Phe)[58] for which comprehensive NMR data in aqueous solution were kindly provided by Peter Schmieder.[59]

The scope of present study is to evaluate the accuracy of the enhanced sampling approach in reproducing structural features and relative state ratios.

METHODS

Structure Preparation. All structures were generated and prepared in Molecular Operation Environment (MOE) 2016.08[60] using the implemented protonate3D tool.[61] For the *cis* and *trans* state of the cyclic pentapeptide cyclo-(Pro-Ser-leu-Asp-Val), starting conformations were constructed based on published backbone torsions using the "Protein Builder" tool in MOE. Starting coordinates of cilengitide were obtained by

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

isolating the macrocycle from the complex X-ray structure (PDB 1L5G).[57] To reduce a possible bias introduced by the starting structure, a geometry optimization in vacuum was performed using the Amber10:EHT force field and default settings. This yielded a starting conformation with a backbone RMSD of 1.0 Å (2.6 Å all atom RMSD) to the crystal structure (Figure S2). For the simulations of the cyclic hexapeptide cyclo-(Arg-Arg-Trp-Trp-Arg-Phe), we started from an unpublished NMR structure kindly provided by Peter Schmieder.[58,59]

Simulation Setup. MD simulations were performed and analyzed using the Amber16 simulation package.[62] We used the program tleap to generate an Amber topology and initial coordinate files.[62] The starting coordinates of the macrocycles were prepared as described above, then a octahedral box of TIP3P waters[63] was added so that every macrocycle atom was at least 12 Å from the box boundary. The three model systems were parametrized using the Amber force field 14SB;[64] additional parameters for N-methylated amino acids were taken from Forcefield_NCAA.[65] A thorough equilibration scheme was applied as described previously[66] to minimize and relax the system. NpT MD simulations were run at atmospheric pressure, by applying the Berendsen barostat,[67] and 300 K using the Langevin thermostat.[68] The SHAKE algorithm[69] was applied to restrain all bonds involving hydrogens, allowing a time step of 2 fs. Productive cMD and aMD simulations of 1 μs length were performed using the GPU implementation of pmemd[70] with a nonbonded cutoff of 8 Å and stored as 500,000 equally spaced snapshots. For cilengitide, 20 aMD simulations of 50 ns each (25,000 frames) were run with randomized starting velocities and combined to a 1 μs trajectory for further analysis. In Figure S1 we illustrate that we are not able to capture significant structural changes with the 1 μs cMD simulation using cyclo-(Pro-Ser-leu-Asp-Val) as a representative example.

All aMD simulations were performed using the dual-boost algorithm implemented in Amber16,[62] thus a bias was applied on the total potential with an additional boost on the dihedral term.[39,71] The parameters determining the boosting were derived as proposed by Pierce et al.[42,43] based on average energies from prior cMD simulations, the number of atoms and residues, and are listed in the Supporting Information.

Trajectory Analysis. Thermodynamic information on the unbiased ensemble is reconstructed from the aMD ensemble with Boltzmann reweighting using Maclaurin series expansion up to the 10th order.[40] For small biasing potentials, which exactly show Gaussian distribution, the cumulant expansion to the second order produces the most accurate results.[40,72] For large biomolecular systems and high boosting potentials, Maclaurin series expansion is the more robust method.[42,43,73] To sample the state transitions in the studied macrocycles, we applied parameters that are expected to achieve high boosting potentials, and thus Maclaurin series expansion was used to estimate the exponential term in the reweighting scheme.

The aMD trajectories are analyzed with cpptraj[74] and in-house scripts. A protocol provided by Miao et al.[40] was applied to perform the Boltzmann reweighting. For each system, the covered conformational space was analyzed by performing principal component analysis (PCA) based on the backbone dihedral angles φ, ψ, and ω as implemented in cpptraj. Likewise, also principal components based on Cartesian coordinates of backbone heavy atoms were calculated (Figure S1). For the studied macrocycles, we found that PCA in Cartesian space leads to a less distinct representation of highly populated states compared to dihedral PCA. Additionally, we characterized the internal motions using the eccentricity ε[75] as a global variable for each snapshot. A value of ε near 1 describes an aspherical compound and 0 indicates perfect globularity. As shown in Figure S3 we find that this metric is able to capture the topological differences between the *cis* and *trans* state of cyclo-(Pro-Ser-leu-Asp-Val). Yet, similar to the Cartesian PCA, the differentiation between densely sampled states is not as distinct as observed in the dihedral PCA space. Therefore, for further investigations we used the representation of the structural ensemble in dihedral space.

The resulting PCA plots of the reweighted structural data projected on the respective eigenvectors were created using Gnuplot 5.0.[76] To study varying structural properties, dihedral angles and backbone RMSD to the bioactive conformation were calculated with cpptraj and used to color-code the projected data in PCA coordinates. Furthermore, reweighted flexibilities were retrieved by calculating dihedral entropies using in-house scripts.[43] Dihedral entropies are an alignment-independent metric to quantify backbone motions of proteins and peptides and have been shown to be applicable for cMD as well as aMD simulations.[43,77]

To simplify the clustering, we benchmarked the minimum number of representative snapshots required to achieve reliable and coherent results. We applied an energetic cutoff based on the boosting potential discarding high-energy conformations. For the studied systems, we found that incorporating 2000 snapshots, that is, 0.4% of the total number of snapshots, significantly reduces the calculation time and increases robustness of cluster analysis compared to using all 500,000 snapshots. We found that including only snapshots with high boosting potentials consistently leads to cluster representatives that are located near the free energy minima in PCA space. Height and distribution of the boosting potentials are strongly varying for each system, hence so does the range of the boosting potential of the top 2000 snapshots. For cyclo-(Pro-Ser-leu-Asp-Val), the incorporated snapshots comprise boosting potentials from 31 to 49 kcal/mol, and for cilengitide, this window ranges from 27 to 42 kcal/mol and for the largest system cyclo-(Arg-Arg-Trp-Trp-Arg-Phe) from 31 to 51 kcal/mol. The average interproton distances and asymmetric errors calculated from the reduced ensemble are consistent with the full ensemble. However, the optimal number of snapshots entirely relies on the analysis objectives and is strongly dependent on the system under investigation.

We conducted a hierarchical agglomerative clustering based on the sine and cosine contributions of the backbone dihedrals using cpptraj to accomplish a thorough structural characterization of representative conformations. The number of clusters was used as a stop criterion for the clustering and was gradually increased until at least one representative structure was found for each free energy minimum in the reweighted PCA.

To compare the aMD ensemble to experimental NOE measurements, we calculated an inverse sixth power ensemble average of the reweighted hydrogen distances, as described in the literature.[78,79] The aMD reweighting procedure was performed using a python toolkit by Miao et. al as described above. An additional distance weighting, that is, the use of r^{-6} averaging is necessary to make aMD and NMR-derived distances comparable.[79] As expected, we observed that the sampled NOE distances are not evenly distributed around the average. To account for this asymmetric distance distribution, we determined asymmetric error bars by separately calculating

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

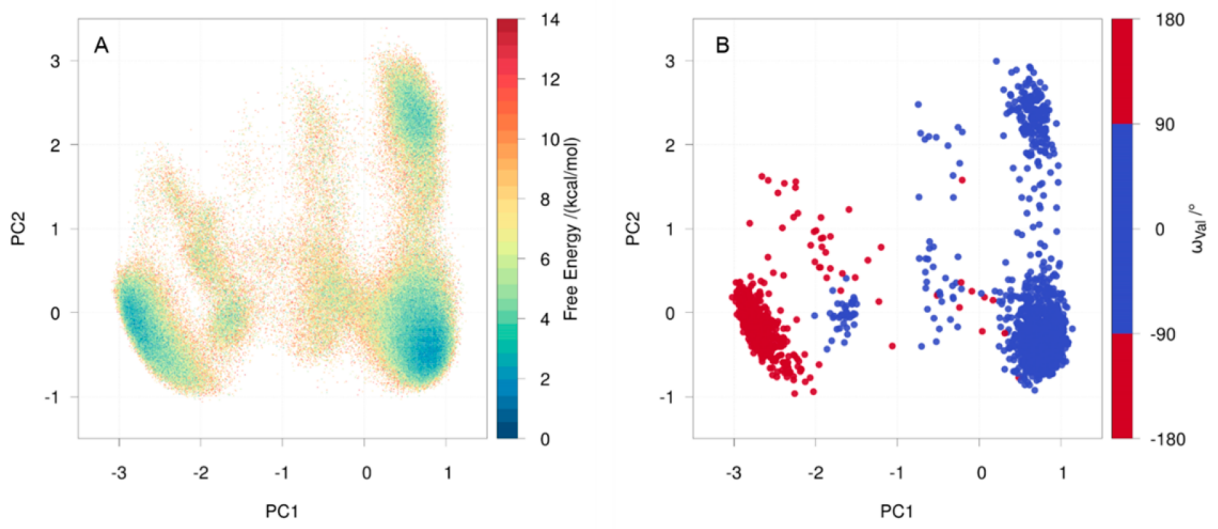


Figure 2. Conformational space of cyclo-(Pro-Ser-leu-Asp-Val). (A) The aMD ensemble was color-coded according to the reweighted free energies and depicted as projection onto the first two PCA eigenvectors. (B) The 2000 most favorable snapshots, color-coded according to their conformational state, were projected onto the same PCA eigenvectors. Structures with $|\omega_{Val\text{-}5}| \leq 90°$ were considered as *cis* (blue) and the remaining structures with $|\omega_{Val\text{-}5}| > 90°$ as *trans* state (red).

#	Atom 1	Atom 2	NMR	lower bound	upper bound	aMD
CIS						
1	Ser HN	Ser Hb	2.6	1.6	3.8	2.5
2	Ser HN	Ser Hb	2.6	1.6	3.8	2.8
3	Ser HN	Val Ha	2.7	1.7	3.7	3.6
4	Ser HN	Pro Ha	3.4	1.9	4.2	3.1
5	Leu HN	Ser Ha	2.1	1.5	3.1	2.2
6	Leu HN	Leu Ha	2.1	1.5	3.1	2.9
7	Asp HN	Leu Ha	2.0	1.5	3.0	2.1
8	Asp HN	Asp Ha	3.4	2.7	4.2	2.9
9	Asp HN	Asp Hb	2.6	1.6	3.9	2.5
10	Asp HN	Asp Hb	2.6	1.6	3.9	2.5
11	Val HN	Asp HN	3.5	2.5	4.5	2.3
12	Val HN	Asp Ha	3.0	2.2	4.1	2.6
13	Val HN	Asp Hb	3.0	1.9	3.7	2.7
14	Val HN	Asp Hb	3.0	1.9	3.7	2.3
15	Val HN	Val Ha	2.5	1.7	3.2	2.8
16	Val HN	Val Hb	2.5	1.5	3.5	2.7
17	Val Ha	Val Hb	2.3	1.5	3.3	2.5
18	Val Ha	Pro Ha	1.9	1.5	2.6	2.1

#	Atom 1	Atom 2	NMR	lower bound	upper bound	aMD
TRANS						
1	Ser HN	Ser Ha	3.0	2.0	4.0	2.8
2	Leu HN	Ser Ha	2.2	1.5	3.2	2.1
3	Asp HN	Leu Ha	2.7	1.7	3.7	2.2
4	Asp HN	Asp Ha	2.5	1.5	3.5	2.9
5	Asp HN	Asp Hb	2.8	1.8	4.2	2.4
6	Asp HN	Asp Hb	2.8	1.8	4.2	2.5
7	Val HN	Asp Ha	2.5	1.5	3.5	3.4
8	Val HN	Asp Hb	2.7	1.7	3.8	2.8
9	Val HN	Asp Hb	2.7	1.7	3.8	2.4
10	Val HN	Val Hb	2.6	1.6	3.6	2.5

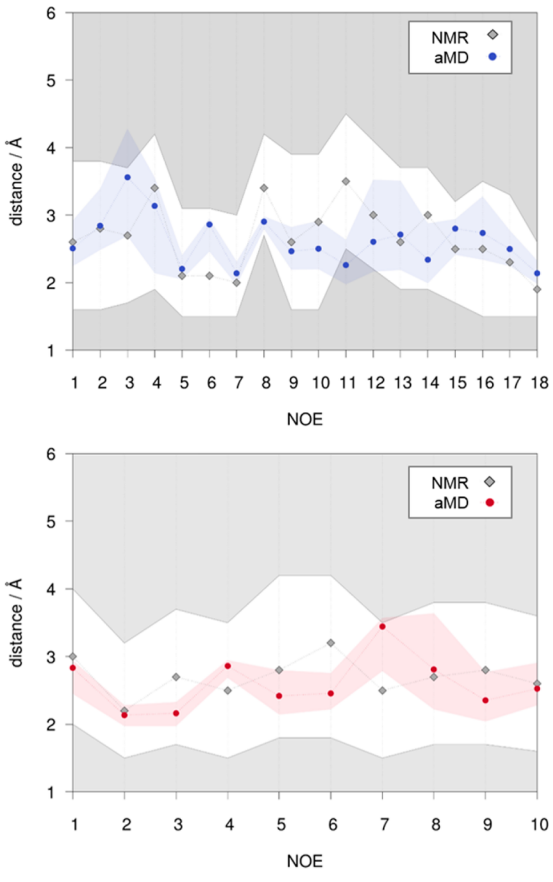


Figure 3. Comparison of interproton distances in Å from NOESY experiments and aMD simulations of the *cis* and *trans* state of cyclo-(Pro-Ser-leu-Asp-Val). The gray diamonds indicate the experimentally estimated average distance. The allowed area within the NMR boundaries is depicted in white. Average distances calculated from the aMD snapshots corresponding to the *cis* and *trans* state with the according asymmetric error bands are shown in color.

the population weighted RMSD from the mean NOE distance for values above and below the mean.

Experimental References. Special care has to be taken when comparing interproton distances from MD simulations and NMR experiments. First of all, the sign of the NOE in NOESY spectra depends on the correlation time of the molecule. Particular unfortunate combinations of field strength and molecular size can even lead to zero or near zero NOEs.

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

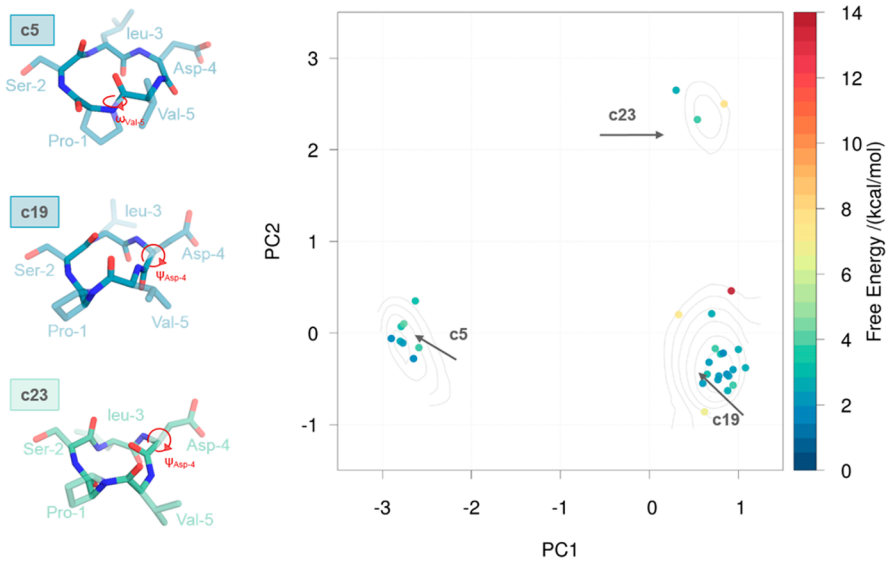


Figure 4. Representative structures of the cyclo-(Pro-Ser-leu-Asp-Val) ensemble. The 30 representative structures are color-coded according to their reweighted free energies and projected onto the PCA eigenvectors. Free energy minima in the PCA space are indicated with gray contour lines from 5.2 to 6.2 kcal/mol in 0.2 kcal/mol steps. Representatives of cluster c23 and c19 were both identified as *cis* clusters differing most prominently in the backbone torsion $\psi_{\text{Asp-4}}$, see Table S2 for further details.

Thus, the measured NOE may be biased toward too long distances or completely vanish even though two protons are close in space.[52,80] Furthermore, experimental NOE signals may overlap and thus cannot be resolved.[52,80] Hence, not all close hydrogen distances observed in MD simulations are expected to be found experimentally, but all experimental NOEs should be present within a given range in our MD simulations.

Cyclo-(Pro-Ser-leu-Asp-Val). A comprehensive study characterized the conformational ensemble of the cyclic pentapeptide cyclo-(Pro-Ser-leu-Asp-Val) using NMR experiments.[54] In the ^{1}H-spectrum published in this work, five slowly interconverting conformational states were identified, but only two of them were sufficiently populated to distinguish them in the NOESY spectrum. Hence, hydrogen distance restraints and populations for two out of the five states were reported, which were identified as *cis*- and *trans*-proline states (Figure 1).

Cilengitide. Structure and activity of the integrin binding peptide cilengitide have been thoroughly characterized in several experimental studies.[55,56,81] The available structural data comprise distance restraints from ROESY experiments in aqueous solution[55] as well as a crystal structure of the cyclic pentapeptide bound to its target integrin (PDB 1L5G)[57] (Figure 1).

Cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). Structural characteristics of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe) were profiled in an extensive NMR study comprising 2D NOESY and ROESY spectra in varying environments.[58] To determine the structure in aqueous solution, a simulated annealing protocol was applied using ROE and NOE distance restraints as well as homonuclear coupling constants. Based on the high number of NOE/ROE violations and high RMSD values, the cyclic hexapeptide was reported as a remarkably flexible macrocycle in water (Figure 1).

RESULTS

Cyclo-(Pro-Ser-leu-Asp-Val). To illustrate the sampled conformational space, we performed a PCA. We found that 36.9% and 14.3% of the total variance of dihedral movements in the simulation were described by PC1 and PC2, respectively. In Figure 2A, the 1 μs aMD trajectory was projected onto the first two principal component vectors (PC1 and PC2) and color-coded according to the Boltzmann reweighted free energy. In the conformational space spanned by PC1 and PC2, we found at least three clearly separated free energy minima located approximately at [−3, −0.5], [0.8, −0.5], and [0.8, 2.5].

As described in the Methods section, we were focusing on the 2000 snapshots with the highest boosting potential, that is, the snapshots with the lowest potential energy. We compared relative state ratios from aMD simulations and NMR experiments by calculating the backbone torsion of Valine-5, $\omega_{\text{Val-5}}$ (highlighted in Figure 1B), as defined by Viles et al.[54] to distinguish between *cis*- and *trans*-proline states. Structures with $|\omega_{\text{Val-5}}| \leq 90°$ were considered as *cis* and the remaining structures with $|\omega_{\text{Val-5}}| > 90°$ as *trans* conformational states. Within the 2000 investigated snapshots, we found 491 structures in the *trans* state and 1509 in the *cis* state. This corresponds to a distribution of 25/75 which agrees strikingly well with the experimentally reported *trans* to *cis* state ratio of 20/80 (Figure 2B). Furthermore, we reweighted the state distribution of the full aMD ensemble, again finding a *trans* to *cis* state ratio of 25/75.

The average interproton distances and the respective asymmetric errors are shown in Figure 3, where we plotted the aMD distances for the *cis* and *trans* ensemble of cyclo-(Pro-Ser-leu-Asp-Val) compared to the published NMR distances and boundaries. For the *trans* structure, 10 NOE restraints were reported, which were all fulfilled by the weighted ensemble average of the aMD distance distributions. The *cis* state was characterized by 18 NOE restraints. In the simulated ensemble, all but one restraint, NOE 11, were fulfilled. The distance between the amide hydrogen of Val-5 and Asp-4 C_α hydrogens was found to be too short in the aMD ensemble compared to the experimental boundaries. However, as discussed in the Methods section, it is possible that NOE-derived distance boundaries underestimate the true distance between two atoms.

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

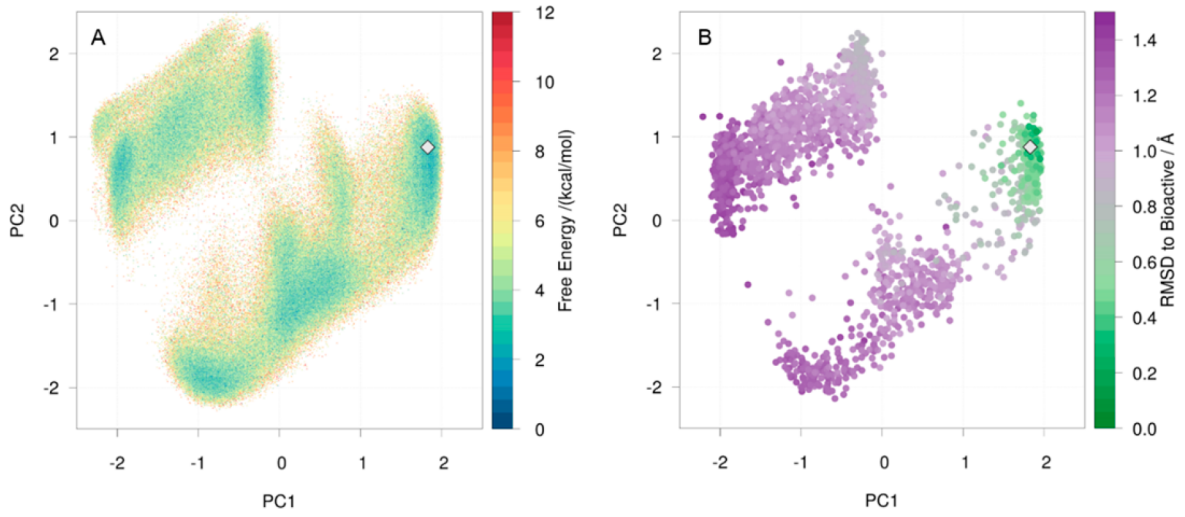


Figure 5. Conformational space of cilengitide. (A) The aMD ensemble was color-coded according to the reweighted free energies and depicted as projection onto the first two PCA eigenvectors. The gray diamond marks the bioactive conformation. (B) The 2000 lowest energy snapshots were projected onto the PCA eigenvectors and color-coded according to the RMSD to the bioactive conformation.

#	Atom 1	Atom 2	upper bound	aMD
1	Arg HN	Arg Ha	2.78	2.30
2	Arg HN	Gly HN	3.74	2.07
3	Arg HN	Val Ha	2.65	2.08
4	Gly HN	Arg Ha	2.94	2.27
5	Asp HN	Asp Ha	2.90	2.88
6	Asp HN	Phe HN	3.68	3.92
7	Phe HN	Asp Ha	2.74	2.20
8	Phe HN	Phe Ha	3.50	2.88
9	Phe HN	Val Ha	5.19	5.00

*

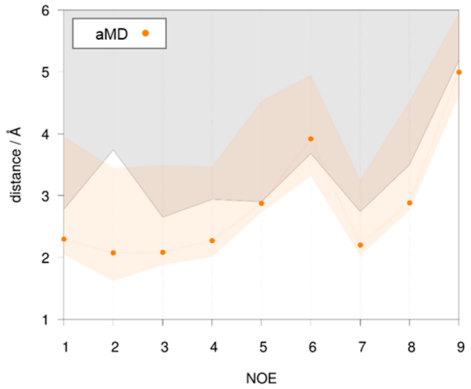


Figure 6. Comparison of interproton distances in Å from NMR experiments and aMD simulations of cilengitide. Upper distance boundaries from ROESY experiments in aqueous solution were used to benchmark the structural accuracy of the aMD ensemble. The white background area indicates distances that fulfill the NMR-derived restraints. Interproton distances calculated from the aMD ensemble of cilengitide with the respective asymmetric errors bands are depicted in orange.

Thus, our results do not necessarily disagree with the experimental data.

For a detailed characterization of the structural features captured with aMD simulations, we clustered the 2000 snapshots with the highest boosting potential—thus lowest energy—resulting in 30 representative structures. In Figure 4 the cluster representatives were projected onto the PCA space. As the color coding in Figure 2B shows, the cluster in the highly populated area on the left-hand side of the PCA around [−3, −0.5] represented structures in the *trans* state. Cluster structures from the two remaining minima on the right-hand side of the PCA plot both represented *cis*-proline states. The structures around [0.8, 0.0] and [0.8, 2.5] could be distinguished by the dihedrals of the amide bond between Asp-4 and Val-5. In structures of cluster c23, representing the least populated minimum, the carbonyl oxygen was oriented toward the center of the ring. In cluster c19 which was found in the center of the most populated minimum around [0.8, 0.0], the same carbonyl oxygen is oriented perpendicular to the ring plane.

In summary, Figure 2 indicated that three to five states were visited in the aMD simulation. The clustering result (Figure 4) further highlighted that at least three clearly distinct conformational states were densely sampled by the aMD ensemble. These strongly sampled areas in the PCA comprised one *trans* and two different *cis* states, which is in good agreement with the available NMR experiments.

Cilengitide. We projected the reweighted structural data of cilengitide on the first two PCA eigenvectors, which comprise 25% and 22% of the total variance (Figure 5A). The resulting free energy surface showed two distinct areas separated by the second eigenvector of the PCA. In each of the two densely sampled areas, several local minima were connected by rather broad pathways. The receptor-bound conformation (gray diamond) was located in the free energy minimum around [1.8, 0.8]. The 2000 conformations with the highest boost are displayed in Figure 5B, color-coded according to their RMSD to the bioactive conformation. We observed that snapshots that were found in the same free energy minimum as the bioactive conformation in PCA space also showed structural characteristics highly similar to the target bound conformation.

The comparison of the calculated average interproton distances with the experimentally derived NOE boundaries showed that 9 out of 10 experimental upper boundaries are

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

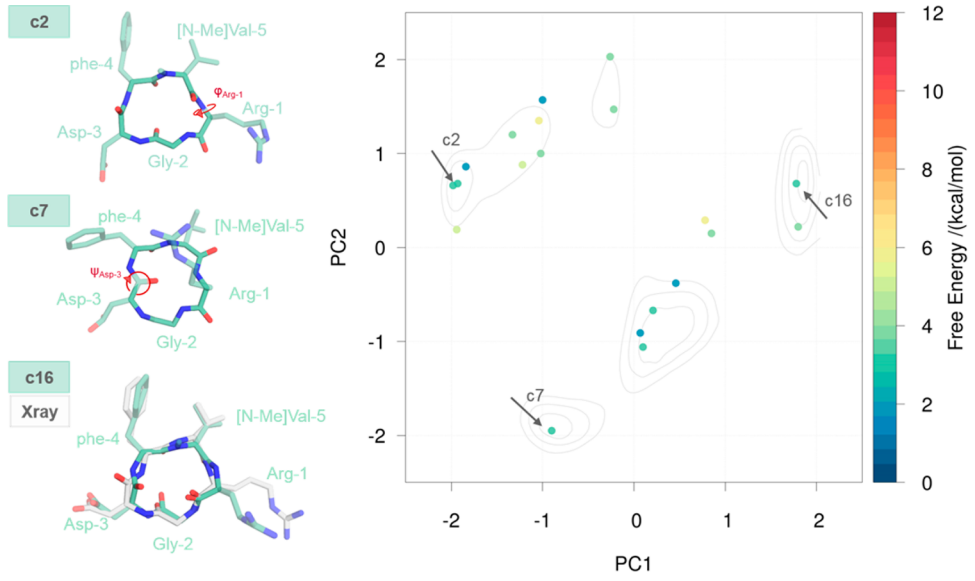


Figure 7. Twenty representative structures of the cilengitide ensemble obtained by cluster analysis. The representative structures were color-coded according to their reweighted free energies and projected onto the PCA eigenvectors. The free energy minima in the PCA space are indicated with gray contour lines from 4.6 to 4.9 kcal/mol in 0.1 kcal/mol steps. The backbone conformation of cluster c16 closely resembled the bioactive conformation. When comparing representatives of other minima with the bioactive conformation (gray diamond), the most prominent differences were reflected by $\varphi_{\text{Arg-1}}$ and $\psi_{\text{Asp-3}}$. See Table S2 for further details.

	Atom 1	Atom 2	lower bound	upper bound	aMD
1	Ser HN	Ser Hb	2.06	3.06	2.85
2	Ser HN	Ser Hb	2.15	3.15	2.40
3	Ser HN	Val Ha	2.13	3.13	2.55
4	Ser HN	Pro Ha	2.13	3.13	2.13
5	Leu HN	Ser Ha	2.24	3.24	2.99
6	Leu HN	Leu Ha	2.22	3.22	2.90
7	Asp HN	Leu Ha	2.02	3.02	2.39
8	Asp HN	Asp Ha	2.12	3.12	2.64
9	Asp HN	Asp Hb	2.02	3.02	2.08
10	Asp HN	Asp Hb	2.03	3.03	2.87
11	Val HN	Asp HN	1.97	2.97	2.21

*

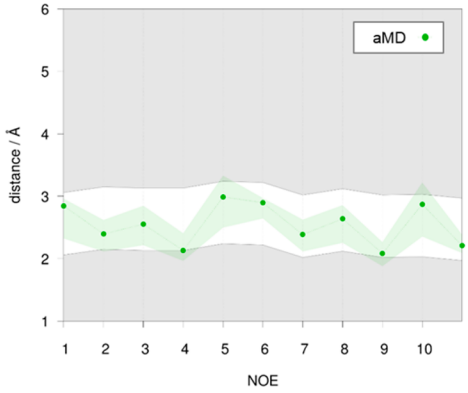


Figure 8. Comparison of interproton distances in Å from ROESY experiments and aMD simulations of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). The white background area indicates distances that fulfill the NMR-derived restraints. Average distances calculated from the aMD ensemble and the according asymmetric error band are depicted in green.

fulfilled by the aMD ensemble. (Figure 6) The outlier, NOE 6, represented the distance between amide hydrogens of Asp-3 and phe-4, which on average was sampled as 0.22 Å too long in the aMD simulation. The violation was rather small and might result from force field inaccuracies, especially in the description of the D-phenylalanine.

Clustering the reduced aMD ensemble, as described earlier, allowed more detailed structural investigations. Figure 7 shows the projection of the resulting cluster representatives onto PC1 and PC2 and selected representative snapshots to illustrate structural features occurring in the individual free energy basins. The representative of cluster c16 represented the same minimum as the bioactive conformation, and the superposition of both structures, shown in Figure 7, highlighted their conformational similarity with a backbone RMSD of 0.6 Å. Further details can be found in Table S3.

Cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). As our third example, we characterized the conformational space of the antimicrobial cyclic hexapeptide cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). In the performed PCA, the first two eigenvectors comprised 33% and 15% of the total variance of dihedral motions. Projection of the 1 μs aMD trajectory onto PCA space yielded a potential energy surface with a large number of minima, indicating a very diverse conformational ensemble (Figure S3). This is in line with NMR structure calculations, where large RMSD values were reported for this macrocycle in aqueous solution.

The average interproton distances, calculated from 2000 snapshots, were all consistent with the NMR-derived boundaries (Figure 8).

To illustrate some of the conformational motifs obtained with the aMD approach, we discretized the ensemble into 15 clusters. Projecting the representative snapshots onto the PCA space, we note that the selected cluster structures represented the minima obtained from the full ensemble (Figure 9).

DISCUSSION

In the present study, we applied aMD simulations as a reliable tool to characterize the structural as well as thermodynamic

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

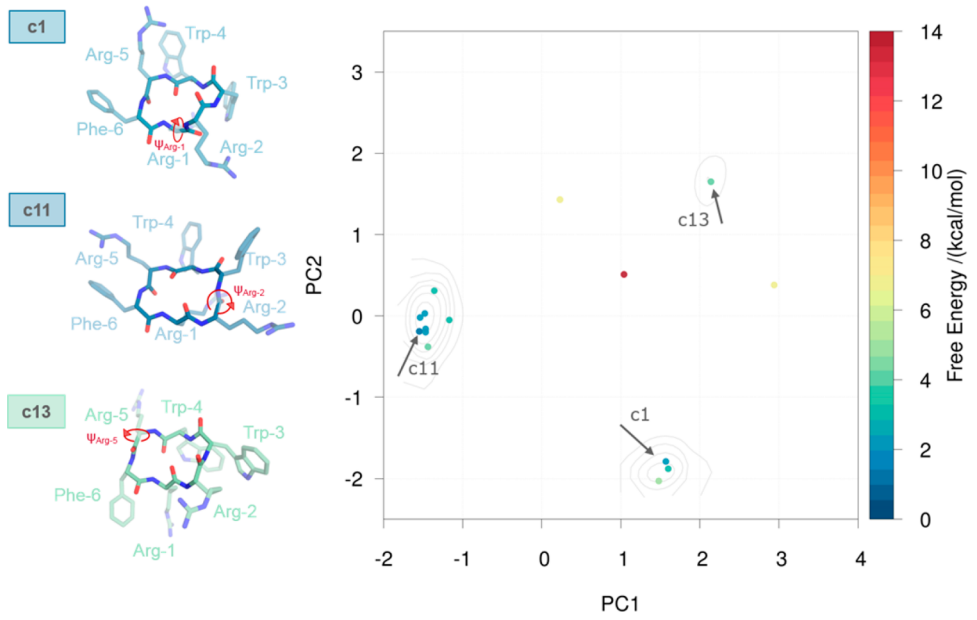


Figure 9. Representative structures of the cyclo-(Arg-Arg-Trp-Trp-Arg-Phe) ensemble obtained by clustering. The representative structures were color-coded according to their reweighted free energies and projected onto the PCA eigenvectors. The free energy minima in the PCA space are indicated with gray contour lines reaching from 6.2 to 7.4 kcal/mol in 0.2 kcal/mol steps. See Table S5 for further details.

properties of peptidic macrocycles. We assessed the performance of the approach for three different cyclic peptides with varying physicochemical properties.

While previous studies reported unsatisfying structural diversity from conventional MD simulations,[14] we showed that the aMD enhanced sampling approach is able to efficiently sample a very diverse ensemble and reliable relative state ratios.

We used PCA as a straightforward tool to illustrate the conformational space covered in the simulations of each cyclopeptide. For the first system, cyclo-(Pro-Ser-leu-Asp-Val), we found at least three clearly delineated free energy minima in the reweighted PCA plot. The representation of the conformational ensemble in PCA space is well in line with NMR experiments on this cyclic pentapeptide, which report clearly distinguishable, slowly interconverting conformational states. The antimicrobial cyclo-(Arg-Arg-Trp-Trp-Arg-Phe) on the other hand shows exceptionally high flexibility in NMR experiments in aqueous solution. The cyclic hexapeptide interconverts relatively fast between different conformational states, as reflected by a high number of NOE violations and large RMSD values in the NMR structure calculations. The PCA from aMD simulations of this system showed a large number of rather broad minima and maxima, indicating a high number of rapidly interchanging conformational states. Compared to the other, less flexible, investigated compounds, the ensemble of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe) also shows a higher structural variance as indicated by the significantly larger axes of the PCA. Thus, PCA of the aMD ensemble reflects the experimentally determined trends in conformational flexibility. To affirm these findings, we further quantified the backbone flexibility of each macrocycle using sum of dihedral entropies[43,77] (Table S4) and again found good agreement with the experimental trends in flexibility.

To assess the structural accuracy of the macrocycle ensemble generated with aMD, we calculated average interproton distances and compared them to distance restraints from NOESY and ROESY experiments. Observables from NMR experiments excel as reference for MD simulations as they provide information on the entire structural ensemble of a compound. For all tested systems, we find reasonable agreement with the NMR restraints. For the highly flexible macrocycle cyclo-(Arg-Arg-Trp-Trp-Arg-Phe), all NMR restraints are fulfilled by the aMD ensemble. Yet, calculation of interproton distances for single representative structures shows large violations, which further highlights that NOE restraints from NMR experiments represent ensemble averages and not necessarily single structural states (Figure S2). Hence, the good agreement with the NMR restraints implies physically meaningful sampling of the conformational space of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe).

For the *cis* state of cyclo-(Pro-Ser-leu-Asp-Val), one distance is found to be on average too short in the aMD ensemble. As described in the Methods section, the possibility of a zero crossing of the NOE cannot be ruled out. Thus, the experimental interproton distance restraint might also be artificially shifted toward higher values, and we do not consider this an alarming result. In the cilengitide ensemble, on the other hand, we find one average interproton distance as being 0.22 Å too long in the aMD ensemble. This outlier clearly indicates a slight discrepancy between aMD and experimental state populations. This discrepancy might derive from simplifications of the force field, for example, D- and L-amino acids being described by the same parameters. Nevertheless, considering the accuracy of the experimental reference paired with the limitation of current force fields, the overall agreement of aMD and NMR is strongly convincing.

Further experimental indications on the structural features of the studied macrocycles were given by $^3J_{\alpha NH}$ coupling constants. These NMR coupling constants can be related to the backbone dihedral angle Θ ($\Theta = |\varphi - 60|$ for L-amino acids and $\Theta = |\varphi + 60|$ for D-amino acids) via the Karplus equation.[54,78] We calculated the reweighted distribution of Θ for the *cis* and *trans* state of cyclo-(Pro-Ser-leu-Asp-Val) (Figure S5). From the published $^3J_{\alpha NH}$ coupling constants, we would

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

expect two significant changes in the respective backbone torsion. On the one hand, the experimental data imply a shift of the Θ(Val) distribution toward larger torsional angles when switching from *cis* to *trans*. For the Θ(Asp) distribution, on the other hand, the NMR measurements indicate a shift toward smaller angles in the *trans* state compared to the *cis* state. As shown in Figure S5, we find good agreement between the distinct structural preferences of *cis* and *trans* states captured with NMR and aMD. However, the agreement between average $^3J_{\alpha NH}$ coupling constants estimated by NMR coupling constants and aMD simulations is rather unsatisfying with a RMSD of 2 Hz. Due to the nature of the Karplus relation, calculation of $^3J_{\alpha NH}$ coupling constants from dihedral data is very sensitive. Small fluctuations of Θ around large angles lead to large deviations in the average $^3J_{\alpha NH}$. Considering the height of the boosting potential which is applied on all torsions and the overall potential, it is not expected to achieve an accurate reproduction of $^3J_{\alpha NH}$ coupling constants. The strength of the aMD approach lies in the efficient sampling of a large portion of the conformational space but in the reweighting step, a large, so-called "energetic noise" is introduced.[41] Consequently, the overall effects of large conformational rearrangements on dihedral distributions are captured. But we observe that the method does not provide the resolution to reweight the subtle differences within the dihedral distributions, which would be needed to accurately estimate the experimental coupling constants. Nevertheless, while the small dissonances in the backbone angle distributions lead to large deviations in the average $^3J_{\alpha NH}$ coupling constants, the overall structures still agree well with experiment, as indicated by interproton distances.

Summarizing, we showed that aMD simulations adequately capture the dynamic and structural properties of peptidic macrocycles in solution. Even more strikingly, we found that also the relative ratio of *trans*- and *cis*-proline states of the integrin binding peptide cyclo-(Pro-Ser-leu-Asp-Val) is coherent with the published NMR data. Despite the known limitations in current force fields and sampling techniques, the presented results strongly indicate that aMD simulations are well suited to structurally and also thermodynamically profile the conformational space of peptidic macrocycles.

CONCLUSION

We evaluated the performance of explicit solvent aMD simulations as conformer generator for peptidic macrocycles. Applying the approach on three different cyclic peptides with varying structural characteristics, we find excellent agreement with experimentally determined structural features.

We showed that aMD simulations are a valid approach to capture a diverse as well as structurally and thermodynamically meaningful conformational ensemble for peptidic macrocycles. The proposed approach allows an extremely thorough characterization of the conformational space and the dynamic behavior of cyclic peptides. Furthermore, for cilengitide we found that also the target-bound conformation is among the frequently sampled states in the aMD simulation. Hence, when the target-bound conformation is already significantly populated in the solution ensemble, aMD identifies the bioactive conformation as a key state, which should be considered in further analysis. So far, the approach is computationally rather expensive. Yet, the calculations can easily be parallelized and further optimized to increase the sampling efficiency.

Altogether, aMD simulations are an excellent technique to perform a comprehensive study on promising macrocyclic compounds. Besides its function as a conformation generator, the aMD approach also allows to estimate thermodynamic properties, such as the flexibility of macrocycles. Furthermore, it is possible to trace changes in the relative ratio of stable conformational states upon modifications within the backbone or functional groups of a macrocyclic compound. The comprehensive computational characterization can be extremely beneficial in macrocycle drug design projects, where synthesis is often rather complicated and costly. Thus, we surmise that aMD simulations hold great potential to significantly enhance the development of macrocyclic drugs.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jcim.8b00097.

Figure S1: Conformational space of cyclo-(Pro-Ser-leu-Asp-Val) captured in 1 μs cMD simulation. Table S1: Applied boosting parameters. Figure S2: Cartesian PCA of cyclo-(Pro Ser-leu-Asp-Val). Figure S2: Comparison to bioactive conformation of cilengitide. Figure S3: Eccentricity of cyclo-(Pro-Ser-leu-Asp-Val). Figure S4: Reweighted distribution of $\omega_{Val\text{-}5}$ in cyclo-(Pro-Ser-leu-Asp-Val). Figure S5: Comparison to bioactive conformation of cilengitide. Table S2: Dihedral angles of selected cluster representatives of cyclo-(Pro Ser-leu-Asp-Val). Table S3: Dihedral angles of selected cluster representatives of cilengitide. Figure S6: PCA of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). Figure S7: Interproton distances of selected cluster representatives of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). Table S4: Sum of dihedral entropies. Table S5: Dihedral angles of selected cluster representatives of cyclo-(Arg-Arg-Trp-Trp-Arg-Phe). Figure S8: Θ dihedral distributions of the *cis* and *trans* states of cyclo-(Pro-Ser-leu-Asp-Val) (PDF)

AUTHOR INFORMATION

Corresponding Authors

*Email: klaus.liedl@uibk.ac.at.

*Email: thomas.fox@boehringer-ingelheim.com.

ORCID

Klaus R. Liedl: 0000-0002-0985-2299

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This work was supported in part by the Austrian Science Fund (FWF) via the grant P30737 "Protein Dynamics and Proteolytic Susceptibility" and in part by Boehringer Ingelheim Pharma GmbH & Co. KG. The authors thank Peter Schmieder for providing experimental data.

REFERENCES

(1) Driggers, E. M.; Hale, S. P.; Lee, J.; Terrett, N. K. The Exploration of Macrocycles for Drug Discovery - an Underexploited Structural Class. *Nat. Rev. Drug Discovery* **2008**, *7*, 608−24.

(2) Villar, E. A.; Beglov, D.; Chennamadhavuni, S.; Porco, J. A., Jr.; Kozakov, D.; Vajda, S.; Whitty, A. How Proteins Bind Macrocycles. *Nat. Chem. Biol.* **2014**, *10*, 723−31.

(3) Over, B.; Matsson, P.; Tyrchan, C.; Artursson, P.; Doak, B. C.; Foley, M. A.; Hilgendorf, C.; Johnston, S. E.; Lee, M. D. t.; Lewis, R. J.; McCarren, P.; Muncipinto, G.; Norinder, U.; Perry, M. W.; Duvall, J. R.; Kihlberg, J. Structural and Conformational Determinants of Macrocycle Cell Permeability. *Nat. Chem. Biol.* **2016**, *12*, 1065−74.
(4) Zorzi, A.; Deyle, K.; Heinis, C. Cyclic Peptide Therapeutics: Past, Present and Future. *Curr. Opin. Chem. Biol.* **2017**, *38*, 24−29.
(5) Hill, T. A.; Shepherd, N. E.; Diness, F.; Fairlie, D. P. Constraining Cyclic Peptides to Mimic Protein Structure Motifs. *Angew. Chem., Int. Ed.* **2014**, *53*, 13020−41.
(6) Nielsen, D. S.; Lohman, R. J.; Hoang, H. N.; Hill, T. A.; Jones, A.; Lucke, A. J.; Fairlie, D. P. Flexibility Versus Rigidity for Orally Bioavailable Cyclic Hexapeptides. *ChemBioChem* **2015**, *16*, 2289−93.
(7) Marsault, E.; Peterson, M. L. Macrocycles Are Great Cycles: Applications, Opportunities, and Challenges of Synthetic Macrocycles in Drug Discovery. *J. Med. Chem.* **2011**, *54*, 1961−2004.
(8) Suarez, D.; Diaz, N. Ligand Strain and Entropic Effects on the Binding of Macrocyclic and Linear Inhibitors: Molecular Modeling of Penicillopepsin Complexes. *J. Chem. Inf. Model.* **2017**, *57*, 2045−55.
(9) Boehr, D. D.; Nussinov, R.; Wright, P. E. The Role of Dynamic Conformational Ensembles in Biomolecular Recognition. *Nat. Chem. Biol.* **2009**, *5*, 789−96.
(10) Carney, D. W.; Schmitz, K. R.; Truong, J. V.; Sauer, R. T.; Sello, J. K. Restriction of the Conformational Dynamics of the Cyclic Acyldepsipeptide Antibiotics Improves Their Antibacterial Activity. *J. Am. Chem. Soc.* **2014**, *136*, 1922−9.
(11) Doak, B. C.; Over, B.; Giordanetto, F.; Kihlberg, J. Oral Druggable Space Beyond the Rule of 5: Insights from Drugs and Clinical Candidates. *Chem. Biol.* **2014**, *21*, 1115−42.
(12) Giordanetto, F.; Kihlberg, J. Macrocyclic Drugs and Clinical Candidates: What Can Medicinal Chemists Learn from Their Properties? *J. Med. Chem.* **2014**, *57*, 278−95.
(13) Bell, I. M.; Gallicchio, S. N.; Abrams, M.; Beese, L. S.; Beshore, D. C.; Bhimnathwala, H.; Bogusky, M. J.; Buser, C. A.; Culberson, J. C.; Davide, J.; Ellis-Hutchings, M.; Fernandes, C.; Gibbs, J. B.; Graham, S. L.; Hamilton, K. A.; Hartman, G. D.; Heimbrook, D. C.; Homnick, C. F.; Huber, H. E.; Huff, J. R.; Kassahun, K.; Koblan, K. S.; Kohl, N. E.; Lobell, R. B.; Lynch, J. J., Jr.; Robinson, R.; Rodrigues, A. D.; Taylor, J. S.; Walsh, E. S.; Williams, T. M.; Zartman, C. B. 3-Aminopyrrolidinone Farnesyltransferase Inhibitors: Design of Macrocyclic Compounds with Improved Pharmacokinetics and Excellent Cell Potency. *J. Med. Chem.* **2002**, *45*, 2388−2409.
(14) Sindhikara, D.; Spronk, S. A.; Day, T.; Borrelli, K.; Cheney, D. L.; Posy, S. L. Improving Accuracy, Diversity, and Speed with Prime Macrocycle Conformational Sampling. *J. Chem. Inf. Model.* **2017**, *57*, 1881−94.
(15) Yudin, A. K. Macrocycles: Lessons from the Distant Past, Recent Developments, and Future Directions. *Chem Sci* **2015**, *6*, 30−49.
(16) Qian, Z.; Dougherty, P. G.; Pei, D. Targeting Intracellular Protein-Protein Interactions with Cell-Permeable Cyclic Peptides. *Curr. Opin. Chem. Biol.* **2017**, *38*, 80−86.
(17) Hickey, J. L.; Zaretsky, S.; St. Denis, M. A.; Kumar Chakka, S.; Morshed, M. M.; Scully, C. C.; Roughton, A. L.; Yudin, A. K. Passive Membrane Permeability of Macrocycles Can Be Controlled by Exocyclic Amide Bonds. *J. Med. Chem.* **2016**, *59*, 5368−5376.
(18) Luther, A.; Moehle, K.; Chevalier, E.; Dale, G.; Obrecht, D. Protein Epitope Mimetic Macrocycles as Biopharmaceuticals. *Curr. Opin. Chem. Biol.* **2017**, *38*, 45−51.
(19) Witek, J.; Keller, B. G.; Blatter, M.; Meissner, A.; Wagner, T.; Riniker, S. Kinetic Models of Cyclosporin a in Polar and Apolar Environments Reveal Multiple Congruent Conformational States. *J. Chem. Inf. Model.* **2016**, *56*, 1547−62.
(20) Rand, A. C.; Leung, S. S.; Eng, H.; Rotter, C. J.; Sharma, R.; Kalgutkar, A. S.; Zhang, Y.; Varma, M. V.; Farley, K. A.; Khunte, B.; Limberakis, C.; Price, D. A.; Liras, S.; Mathiowetz, A. M.; Jacobson, M. P.; Lokey, R. S. Optimizing Pk Properties of Cyclic Peptides: The Effect of Side Chain Substitutions on Permeability and Clearance. *MedChemComm* **2012**, *3*, 1282−89.
(21) el Tayar, N.; Mark, A. E.; Vallat, P.; Brunne, R. M.; Testa, B.; van Gunsteren, W. F. Solvent-Dependent Conformation and Hydrogen-Bonding Capacity of Cyclosporin A: Evidence from Partition Coefficients and Molecular Dynamics Simulations. *J. Med. Chem.* **1993**, *36*, 3757−64.
(22) Fouche, M.; Schafer, M.; Berghausen, J.; Desrayaud, S.; Blatter, M.; Piechon, P.; Dix, I.; Martin Garcia, A.; Roth, H. J. Design and Development of a Cyclic Decapeptide Scaffold with Suitable Properties for Bioavailability and Oral Exposure. *ChemMedChem* **2016**, *11*, 1048−59.
(23) Hawkins, P. C. D. Conformation Generation: The State of the Art. *J. Chem. Inf. Model.* **2017**, *57*, 1747−56.
(24) Friedrich, N. O.; de Bruyn Kops, C.; Flachsenberg, F.; Sommer, K.; Rarey, M.; Kirchmair, J. Benchmarking Commercial Conformer Ensemble Generators. *J. Chem. Inf. Model.* **2017**, *57*, 2719−28.
(25) Glas, A.; Wamhoff, E. C.; Kruger, D. M.; Rademacher, C.; Grossmann, T. N. Increased Conformational Flexibility of a Macrocycle-Receptor Complex Contributes to Reduced Dissociation Rates. *Chem. - Eur. J.* **2017**, *23*, 16157−61.
(26) Chen, I. J.; Foloppe, N. Tackling the Conformational Sampling of Larger Flexible Compounds and Macrocycles in Pharmacology and Drug Discovery. *Bioorg. Med. Chem.* **2013**, *21*, 7898−920.
(27) Bonnet, P.; Agrafiotis, D. K.; Zhu, F.; Martin, E. Conformational Analysis of Macrocycles: Finding What Common Search Methods Miss. *J. Chem. Inf. Model.* **2009**, *49*, 2242−59.
(28) Watts, K. S.; Dalal, P.; Tebben, A. J.; Cheney, D. L.; Shelley, J. C. Macrocycle Conformational Sampling with Macromodel. *J. Chem. Inf. Model.* **2014**, *54*, 2680−96.
(29) Labute, P. Lowmodemd−Implicit Low-Mode Velocity Filtering Applied to Conformational Search of Macrocycles and Protein Loops. *J. Chem. Inf. Model.* **2010**, *50*, 792−800.
(30) Coutsias, E. A.; Lexa, K. W.; Wester, M. J.; Pollock, S. N.; Jacobson, M. P. Exhaustive Conformational Sampling of Complex Fused Ring Macrocycles Using Inverse Kinematics. *J. Chem. Theory Comput.* **2016**, *12*, 4674−87.
(31) Geng, H.; Jiang, F.; Wu, Y. D. Accurate Structure Prediction and Conformational Analysis of Cyclic Peptides with Residue-Specific Force Fields. *J. Phys. Chem. Lett.* **2016**, *7*, 1805−10.
(32) Hosseinzadeh, P.; Bhardwaj, G.; Mulligan, V. K.; Shortridge, M. D.; Craven, T. W.; Pardo-Avila, F.; Rettie, S. A.; Kim, D. E.; Silva, D. A.; Ibrahim, Y. M.; Webb, I. K.; Cort, J. R.; Adkins, J. N.; Varani, G.; Baker, D. Comprehensive Computational Design of Ordered Peptide Macrocycles. *Science* **2017**, *358*, 1461−66.
(33) Laio, A.; Gervasio, F. L. Metadynamics: A Method to Simulate Rare Events and Reconstruct the Free Energy in Biophysics, Chemistry and Material Science. *Rep. Prog. Phys.* **2008**, *71*, 126601.
(34) Gil-Ley, A.; Bussi, G. Enhanced Conformational Sampling Using Replica Exchange with Collective-Variable Tempering. *J. Chem. Theory Comput.* **2015**, *11*, 1077−85.
(35) Sugita, Y.; Okamoto, Y. Replica-Exchange Molecular Dynamics Method for Protein Folding. *Chem. Phys. Lett.* **1999**, *314*, 141−51.
(36) Slough, D. P.; Yu, H.; McHugh, S. M.; Lin, Y. S. Toward Accurately Modeling N-Methylated Cyclic Peptides. *Phys. Chem. Chem. Phys.* **2017**, *19*, 5377−88.
(37) McHugh, S. M.; Rogers, J. R.; Yu, H.; Lin, Y. S. Insights into How Cyclic Peptides Switch Conformations. *J. Chem. Theory Comput.* **2016**, *12*, 2480−8.
(38) Wakefield, A. E.; Wuest, W. M.; Voelz, V. A. Molecular Simulation of Conformational Pre-Organization in Cyclic Rgd Peptides. *J. Chem. Inf. Model.* **2015**, *55*, 806−13.
(39) Hamelberg, D.; Mongan, J.; McCammon, J. A. Accelerated Molecular Dynamics: A Promising and Efficient Simulation Method for Biomolecules. *J. Chem. Phys.* **2004**, *120*, 11919−29.
(40) Miao, Y.; Sinko, W.; Pierce, L.; Bucher, D.; Walker, R. C.; McCammon, J. A. Improved Reweighting of Accelerated Molecular Dynamics Simulations for Free Energy Calculation. *J. Chem. Theory Comput.* **2014**, *10*, 2677−89.
(41) Shen, T.; Hamelberg, D. A Statistical Analysis of the Precision of Reweighting-Based Simulations. *J. Chem. Phys.* **2008**, *129*, 034103.

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992

(42) Pierce, L. C.; Salomon-Ferrer, R.; de Oliveira, C. A. F.; McCammon, J. A.; Walker, R. C. Routine Access to Millisecond Time Scale Events with Accelerated Molecular Dynamics. *J. Chem. Theory Comput.* **2012**, *8*, 2997−3002.
(43) Kamenik, A. S.; Kahler, U.; Fuchs, J. E.; Liedl, K. R. Localization of Millisecond Dynamics: Dihedral Entropy from Accelerated Md. *J. Chem. Theory Comput.* **2016**, *12*, 3449−55.
(44) Miao, Y.; Nichols, S. E.; McCammon, J. A. Free Energy Landscape of G-Protein Coupled Receptors, Explored by Accelerated Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2014**, *16*, 6398−406.
(45) Kappel, K.; Miao, Y.; McCammon, J. A. Accelerated Molecular Dynamics Simulations of Ligand Binding to a Muscarinic G-Protein-Coupled Receptor. *Q. Rev. Biophys.* **2015**, *48*, 479−87.
(46) Fuglestad, B.; Gasper, P. M.; Tonelli, M.; McCammon, J. A.; Markwick, P. R.; Komives, E. A. The Dynamic Structure of Thrombin in Solution. *Biophys. J.* **2012**, *103*, 79−88.
(47) Markwick, P. R.; McCammon, J. A. Studying Functional Dynamics in Bio-Molecules Using Accelerated Molecular Dynamics. *Phys. Chem. Chem. Phys.* **2011**, *13*, 20053−65.
(48) Lopez-Martinez, C.; Flores-Morales, P.; Cruz, M.; Gonzalez, T.; Feliz, M.; Diez, A.; Campanera, J. M. Proline Cis-Trans Isomerization and Its Implications for the Dimerization of Analogues of Cyclopeptide Stylostatin 1: A Combined Computational and Experimental Study. *Phys. Chem. Chem. Phys.* **2016**, *18*, 12755−67.
(49) Caldwell, J. W.; Kollman, P. A. Structure and Properties of Neat Liquids Using Nonadditive Molecular-Dynamics - Water, Methanol, and N-Methylacetamide. *J. Phys. Chem.* **1995**, *99*, 6208−19.
(50) Fox, T.; Kollman, P. A. Application of the Resp Methodology in the Parametrization of Organic Solvents. *J. Phys. Chem. B* **1998**, *102*, 8070−79.
(51) Grabuleda, X.; Jaime, C.; Kollman, P. A. Molecular Dynamics Simulation Studies of Liquid Acetonitrile: New Six-Site Model. *J. Comput. Chem.* **2000**, *21*, 901−08.
(52) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; Wiley: New York, 1986.
(53) We adopt the usual notation of L-amino acids with capital first letter and D-amino acids with all lower case.
(54) Viles, J. H.; Mitchell, J. B.; Gough, S. L.; Doyle, P. M.; Harris, C. J.; Sadler, P. J.; Thornton, J. M. Multiple Solution Conformations of the Integrin-Binding Cyclic Pentapeptide Cyclo(-Ser-D-Leu-Asp-Val-Pro-). Analysis of the (Phi, Psi) Space Available to Cyclic Pentapeptides. *Eur. J. Biochem.* **1996**, *242*, 352−62.
(55) Marelli, U. K.; Frank, A. O.; Wahl, B.; La Pietra, V.; Novellino, E.; Marinelli, L.; Herdtweck, E.; Groll, M.; Kessler, H. Receptor-Bound Conformation of Cilengitide Better Represented by Its Solution-State Structure Than the Solid-State Structure. *Chem. - Eur. J.* **2014**, *20*, 14201−6.
(56) Moruno, C. M. M. M.; Doedens, L.; Frank, A. O.; Kessler, H. Enhancement of Receptor Selectivity of Cilengitide by Multiple N-Methylation. *J. Pept. Sci.* **2010**, *16*, 45−46.
(57) Xiong, J.-P.; Stehle, T.; Zhang, R.; Joachimiak, A.; Frech, M.; Goodman, S. L.; Arnaout, M. A. Crystal Structure of the Extracellular Segment of Integrin Avβ3 in Complex with an Arg-Gly-Asp Ligand. *Science* **2002**, *296*, 151.
(58) Appelt, C.; Wessolowski, A.; Soderhall, J. A.; Dathe, M.; Schmieder, P. Structure of the Antimicrobial, Cationic Hexapeptide Cyclo(Rrwwrf) and Its Analogues in Solution and Bound to Detergent Micelles. *ChemBioChem* **2005**, *6*, 1654−62.
(59) Personal communication with Schmieder, P., 2016.
(60) *Molecular Operating Environment (Moe), 2014.09*; Chemical Computing Group Inc.: Montreal, QC, Canada, 2018.
(61) Labute, P. Protonate3d: Assignment of Ionization States and Hydrogen Coordinates to Macromolecular Structures. *Proteins: Struct., Funct., Genet.* **2009**, *75*, 187−205.
(62) Case, D. A.; Betz, R. M.; Botello-Smith, W.; Cerutti, D. S.; Cheatham, III, T. E.; Darden, T. A.; Duke, R. E.; Giese, T. J.; Gohlke, H.; Goetz, A. W.; Homeyer, N.; Izadi, S.; Janowski, P.; Kaus, J.; Kovalenko, A.; Lee, T. S.; LeGrand, S.; Li, P.; Lin, C.; Luchko, T.; Luo, R.; Madej, B.; Mermelstein, D.; Merz, K. M.; Monard, G.; Nguyen, H.; Nguyen, H. T.; Omelyan, I.; Onufriev, A.; Roe, D. R.; Roitberg, A.; Sagui, C.; Simmerling, C. L.; Swails, J.; Walker, R. C.; Wang, J.; Wolf, R. M.; Wu, X.; Xiao, L.; York, D. M.; Kollman, P. A. *Amber 2016*; University of California: San Francisco, CA, 2016.
(63) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. Comparison of Simple Potential Functions for Simulating Liquid Water. *J. Chem. Phys.* **1983**, *79*, 926−35.
(64) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. Ff14sb: Improving the Accuracy of Protein Side Chain and Backbone Parameters from Ff99sb. *J. Chem. Theory Comput.* **2015**, *11*, 3696−713.
(65) Khoury, G. A.; Smadbeck, J.; Tamamis, P.; Vandris, A. C.; Kieslich, C. A.; Floudas, C. A. Forcefield_Ncaa: Ab Initio Charge Parameters to Aid in the Discovery and Design of Therapeutic Proteins and Peptides with Unnatural Amino Acids and Their Application to Complement Inhibitors of the Compstatin Family. *ACS Synth. Biol.* **2014**, *3*, 855−69.
(66) Wallnoefer, H. G.; Liedl, K. R.; Fox, T. A Challenging System: Free Energy Prediction for Factor Xa. *J. Comput. Chem.* **2011**, *32*, 1743−52.
(67) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; Dinola, A.; Haak, J. R. Molecular-Dynamics with Coupling to an External Bath. *J. Chem. Phys.* **1984**, *81*, 3684−90.
(68) Adelman, S. A.; Doll, J. D. Generalized Langevin Equation Approach for Atom-Solid-Surface Scattering - General Formulation for Classical Scattering Off Harmonic Solids. *J. Chem. Phys.* **1976**, *64*, 2375−88.
(69) Ciccotti, G.; Ryckaert, J. P. Molecular-Dynamics Simulation of Rigid Molecules. *Comput. Phys. Rep.* **1986**, *4*, 346−392.
(70) Salomon-Ferrer, R.; Gotz, A. W.; Poole, D.; Le Grand, S.; Walker, R. C. Routine Microsecond Molecular Dynamics Simulations with Amber on Gpus. 2. Explicit Solvent Particle Mesh Ewald. *J. Chem. Theory Comput.* **2013**, *9*, 3878−88.
(71) Hamelberg, D.; de Oliveira, C. A.; McCammon, J. A. Sampling of Slow Diffusive Conformational Transitions with Accelerated Molecular Dynamics. *J. Chem. Phys.* **2007**, *127*, 155102.
(72) Jing, Z.; Sun, H. A Comment on the Reweighting Method for Accelerated Molecular Dynamics Simulations. *J. Chem. Theory Comput.* **2015**, *11*, 2395−7.
(73) Thomas, J. R.; Gedeon, P. C.; Grant, B. J.; Madura, J. D. Leut Conformational Sampling Utilizing Accelerated Molecular Dynamics and Principal Component Analysis. *Biophys. J.* **2012**, *103*, L1−3.
(74) Roe, D. R.; Cheatham, T. E. 3rd, Ptraj and Cpptraj: Software for Processing and Analysis of Molecular Dynamics Trajectory Data. *J. Chem. Theory Comput.* **2013**, *9*, 3084−95.
(75) Arteca Gustavo, A. Molecular Shape Descriptors. *Reviews in Computational Chemistry*, John Wiley & Sons, Inc.: Hoboken, NJ, 2007.
(76) Williams, T.; Kelley, C. *Gnuplot 5.0, an Interactive Plotting Program*, 2017.
(77) Fuchs, J. E.; Huber, R. G.; Waldner, B. J.; Kahler, U.; von Grafenstein, S.; Kramer, C.; Liedl, K. R. Dynamics Govern Specificity of a Protein-Protein Interface: Substrate Recognition by Thrombin. *PLoS One* **2015**, *10*, e0140713.
(78) Daura, X.; Antes, I.; van Gunsteren, W. F.; Thiel, W.; Mark, A. E. The Effect of Motional Averaging on the Calculation of Nmr-Derived Structural Properties. *Proteins: Struct., Funct., Genet.* **1999**, *36*, 542−55.
(79) Zagrovic, B.; van Gunsteren, W. F. Comparing Atomistic Simulation Data with the Nmr Experiment: How Much Can Noes Actually Tell Us? *Proteins: Struct., Funct., Genet.* **2006**, *63*, 210−8.
(80) Hinds, M. G.; Norton, R. S. Nmr Spectroscopy of Peptides and Proteins. Practical Considerations. *Mol. Biotechnol.* **1997**, *7*, 315−31.
(81) Verma, H.; Khatri, B.; Chakraborti, S.; Chatterjee, J. Increasing the Bioactive Space of Peptide Macrocycles by Thioamide Substitution. *Chem Sci* **2018**, *9*, 2443−51.

DOI: 10.1021/acs.jcim.8b00097
J. Chem. Inf. Model. 2018, 58, 982−992